Exhibit 4

Valuation of Patented Product Features

Greg M. Allenby *Ohio State University*
Jeff Brazell *The Modellers*
John R. Howell *Pennsylvania State University*
Peter E. Rossi *University of California, Los Angeles*

Ultimately, patents have value to the extent to which the product features enabled by the patents have economic value in the marketplace. Products that are enhanced by inclusion of patented features should generate incremental profits. Incremental profits can be assessed by considering demand for products with patented features and contrasting that demand with demand for the same product without the patented features. Profit calculations must be based on valid estimates of demand as well as assumptions about how competitive forces affect demand via computation of market equilibria. A conjoint survey can be used to estimate demand. Recently, conjoint methods have been applied in the patent setting, but the measures of value used are purely demand based and do not involve equilibrium profit calculations. We illustrate our method using the market for digital cameras and show that current methods can overstate the value of a patent.

1. Introduction

Many patents derive their primary value from the product features that are enabled by practicing the patent. For example, consider a patent that can be construed to cover the so-called smart gestures of smartphone operating systems[1] (such as various ways of navigating and zooming in and out using finger swipes). This patent has value only to the extent that this feature is valued by consumers and not achievable by other methods with similar functionality that do not infringe the patent. This is nothing more than an application of the principle

Rossi would like to acknowledge the Collins Chair, Anderson School of Management, University of California, Los Angeles, for research funding. Allenby thanks the Fisher College of Business at Ohio State University for generous research support.

[1] This example is from recent patent litigation between Apple and Samsung (*Apple Electronics Co. Ltd. v. Samsung Electronics Co. Ltd.*, No. 11-CV-1846 [N.D. Cal. December 2, 2011]).

[*Journal of Law and Economics*, vol. 57 (August 2014)]

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 89 of 122

that demand elasticity (and hence market power) is driven by the extent to which competing products are substitutable for the focal product.

Clearly, estimates of the demand or consumer valuation of product features are a vital part of the problem of patent valuation. However, our perspective is that the economic valuation of a patent does not end with a calibration of consumer demand for product features. The economic value of a patent should be based on the incremental profits that can accrue to firms that enhance their products with the features enabled by the patents.[2] These incremental profits are based not just on consumer valuation of features (demand) but also on cost and competitive factors (supply). Therefore, incremental-profit calculations should be based on equilibrium profits. That is to say, we should consider equilibrium in the industry prior to addition of the product features enabled by the patent and then recompute the industry equilibrium after products are enhanced with the features covered by the patent. Incremental equilibrium profit calculations can also provide an economically coherent basis for either reasonable-royalty or lost-profit calculations in the damages portion of patent litigation.

Equilibrium profit calculations are accurate only if we can estimate the industry demand system, measure the marginal cost of feature enhancement, and correctly characterize the nature and extent of competition in the industry. Observational data on demand for products in the marketplace are often inadequate to estimate the demand for patented features, as there is no information on sales of products in the counterfactual case in which the patented features are absent from the industry. In addition, there may be inadequate price variation or concerns about price endogeneity that render observational data uninformative with respect to price sensitivity. Instead, a conjoint survey can be used to calibrate demand. A conjoint survey can be viewed as a market simulator in which consumers are asked to choose among hypothetical products whose features (including price) are varied in a randomized design.

However, a conjoint survey, in and of itself, is not adequate to form the basis for equilibrium firm profit calculations. Not only must we calibrate demand for products, but we must also compute industry equilibria. This requires measures of costs, a demand system not only for the focal product but also for the major competing products, and an equilibrium concept. We argue that a differentiated-products demand system and a Nash equilibrium pricing game provide a reasonable basis for equilibrium calculations in many patent situations.

The practice of patent valuation has typically not involved rigorous quantitative methods or the use of survey data. Instead, judgmental methods are often used without any compelling economic framework. Recently, however, survey-based methods of calibrating demand systems have begun to assume a role in some damage analyses (see, for example, *Apple Electronics Co. Ltd. v. Samsung Elec-*

[2] To be precise, the economic value of a patent is the highest value of the patent in use. That is, the value is determined by the largest incremental profits that could accrue to the firm that adds the product feature or feature enhancement afforded by the patent.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 90 of 122

tronics Co. Ltd., No. 11-CV-1846 [N.D. Cal. December 2, 2011] and *Microsoft Corp. v. Motorola Inc.*, 696 F.3d 872 [9th Cir. 2012]). In this early work, only pure demand-based measures of product feature value have been used. For example, it is possible to compute the change in average willingness to pay (WTP) as a patented feature is added to a product. A pure demand-based measure, WTP is only indirectly related to incremental firm profits. Similarly, the so-called willingness to buy (WTB) analysis of the gains and/or losses in market share attributed to patented product features does not consider price response to the introduction of new products with the patented features.

In summary, we advocate using equilibrium outcomes (both price and shares) to determine the incremental economic profits that would accrue to a firm as a product is enhanced by patented features as a way of computing the economic value of a patent. To compute equilibrium outcomes, we will have to make assumptions about cost, the nature of competition, and the set of competitive offers. Conjoint studies have to be designed with this in mind. In particular, greater care to include an appropriate set of competitive brands, handle the outside option appropriately, and estimate price sensitivity precisely must be exercised. We illustrate our method using a conjoint survey for digital cameras and the addition of a swivel-screen display as the object of the valuation exercise. We show that other methods advocated in recent patent litigation (such as WTP and WTB) can overstate the value of a patent and inflate damages estimates.

2. Economic Valuation of Patents and an Economic Basis for Patent Damages

The primary source of value for many patents is to enable a firm to enhance a product with differentiating features. Economists who have studied innovation in products have emphasized the increase in consumer surplus afforded by the product innovations (see, for example, Trajtenberg 1989; Petrin 2002). While a social welfare point of view is certainly relevant for policy issues such as whether patents should exist in the first place, it is our view that the valuation of patents in the marketplace and the losses incurred from patent infringement are based on firm profitability. What a firm is willing to pay for a patent should be determined, at least in part, by the incremental profits that the features enabled by the patent can create. Similarly, the license fees or royalties charged for use of a patent should depend on the flow of incremental profits to the firm practicing the patent under license.

To compute the incremental profits afforded by a patented feature, we must consider two states of the world. First, we consider a world without the focal feature. In this world, no firm will have access to the feature (or feature enhancement). In this counterfactual world, firms would set prices in an industry equilibrium and obtain profits based not only on demand for the nonenhanced products but also on costs and the degree of competition. The value of the patent

is calculated by adding the patented feature to one or more existing products and recomputing the industry equilibrium. If the feature creates greater demand (increased consumer surplus or WTP), then the firm with the enhanced products may be able to earn higher profits than in the counterfactual world. However, we do not expect that industry equilibrium prices will remain the same in the new world with the enhanced products. Typically, firms whose products do not have the enhanced features will face stiffer competition for customers. In response, these firms will lower their prices, which results in a new industry equilibrium. It is the comparison of the focal firm's profits in the new world with the enhanced product to the firms' profits in the counterfactual world for the product without the feature that we advocate as a reasonable measure of the economic value of the patent.

Patents, by their very nature, can have value only in oligopolistic industries with differentiated products. In industries (such as consumer electronics) with differentiated products and a relatively modest number of competitors, firms can earn rents on the patented features. In perfectly competitive industries or in the case of homogenous products, firms cannot generally earn rents on patented features. This means that we must employ a model of differentiated product demand along with an appropriate equilibrium concept for oligopolistic industries. We employ a characteristics-based discrete-choice model with heterogeneous preferences (see, for example, Berry, Levinsohn, and Pakes 1995), which has now become standard for models of the demand for differentiated products. We compute a static Nash price equilibrium, conditional on a set of competing products defined by points in the characteristics space.

Patent litigation has become a high-stakes game in which large damages claims are made for patents involving consumer products such as gaming consoles, smartphones, and tablet computers. The patents at stake in this litigation involve only a small subset of the features of these electronic devices. Therefore, the question of how to apportion value over a very large set of features is important. Conjoint surveys are beginning to be employed in order to determine what portion of demand or total consumer surplus originates in the patented features. While this is certainly an important first step toward valuing the features, it is not a complete solution to the damages problem. If a firm seeks damages from patent infringement, those damages are almost certainly in the form of either lost profits or license/royalty fees. If the firm that owns the patent also practices the patent, then a lost-profits analysis is appropriate. Only that portion of profits that can be ascribed to the patented features should form the basis of a lost-profits analysis. For example, in the case of Apple suing Samsung over infringement of so-called smart-gesture patents, Apple practices the patent, as the iPhone embodies smart-gesture features. Apple contends that Samsung has infringed on these patents and therefore reduced Apple's profits from what they would have been in the absence of infringement. This requires calculation of incremental profit that compares Apple's profits in the counterfactual situation in which Samsung's smartphone products do not contain features enabled by the patents

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 92 of 122

with Apple's profits earned in the marketplace in which Samsung's products have smart-gesture features. Computation of the lost profits requires that we solve for two industry equilibria corresponding to the cases in which smart-gesture features are present or absent from Samsung's products.

In situations in which the plaintiff firm does not practice the patents under litigation, it is common to seek damages in the form of a license fee or reasonable-royalty analysis. The idea is that there should have been a negotiation between the owner of the patent and the firms accused of infringement to establish a license payment or royalty rate. Other than examining other market transactions involving licensing similar patents, there is no rigorous approach to establishing a license fee. In principle, the basis for this hypothetical negotiation should be the value of practicing the patent on the part of the licensee. That is, if the licensee incorporates the features enabled by the patent at issue, then this should generate a stream of incremental profits from the sales of the enhanced products. These incremental profits stem from a price premium that the licensee can charge for the enhanced product or from increased sales or both. The fundamental point is that these incremental profits should form the basis of the negotiation. In other words, the pie that is being split between the licensor and licensee is made from the incremental profits. Again, the methods advocated here can be used to compute the incremental profits. This should be the basis for the science of a reasonable-royalty analysis. The art continues to be answering the question of how the incremental profits are split between the licensor and licensee. Presumably, this stems from the relative negotiating power of the two parties. We are not contributing to the question of how to establish a splitting rule, but, rather, we provide one way of establishing the profit base to be split.

We now review how a conjoint survey method can form the basis of a valid demand system and, coupled with competitive assumptions, provide a way of computing incremental profits.

3. Defining and Measuring Demand for Product Features

Clearly, any valid paradigm for patent valuation must measure demand in the relevant market. The demand schedule facing any firm's product can be constructed from the distribution of preferences across the population of consumers. We take a characteristics view of products; that is, products are simply bundles of features or characteristics. Firms pick a particular bundle of features, which can be thought of as a point in the space of possible product configurations. The competitive structure of the industry is defined not only by the number of competing products but also by the positions of those products in the characteristics or product feature space. Given a set of products with associated demand, a competitive equilibrium can be computed, which enables computation of the incremental profits flowing to the firm offering a product with patented features.

We use a standard model for the demand for differentiated products. Con-

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 93 of 122

sumers choose only one product at any one purchase occasion and may elect to not purchase any product. A firm introduces a product into the marketplace by announcing a particular configuration of product features or characteristics and a price. The demand schedule facing this firm is the expected sales of the product given the price, configuration of features, and distribution of preferences across consumers. For example, a digital camera product might consist of a brand name (for example, Sony), a particular resolution (measured in megapixels), a particular level of zoom performance, a style of display (swivel screen), and so forth. The firm announces a price for this product. Consumers evaluate this product and compare it with other products in the marketplace. Each consumer makes a choice, and these choices add up to total demand for the product. Each consumer is represented by a specific utility function that is specified by a set of utility weights on the product characteristics. A choice model can predict the probability that a consumer will pick a specific product from a set of products on the basis of the characteristics of each product and the weights that represent that consumer's preferences for features.

3.1. *The Standard Choice Model for Differentiated Product Demand*

The consumer faces a marketplace that consists of J products each specified by a characteristics vector $\boldsymbol{x}_j$ and a price p_j. If there are k characteristics, then the characteristics vector identifies the specific values of each of the k characteristics. In our example, $\boldsymbol{x}_j =$ (Sony, 16 mp, 10 × zoom, without swivel screen). Many characteristics have only a discrete set of possibly unordered values. For these characteristics (like the presence or absence of a swivel-screen display), we introduce a set of binary indicator or dummy variables into the characteristics vector. For example, if there are four major brands in the marketplace, then four indicator (zero/one) variables[3] would be included in the characteristics vector.

The utility derived from the purchase and use of any product is modeled via what has become a standard random utility model (McFadden 1983) for choice applications. The observed characteristics and price enter the utility for the jth brand in a linear fashion. An error term is introduced to account for unobserved factors that influence choice. The error term is often interpreted as representing unmeasured characteristics that influence choice-specific utility. In the equation

$$u_j = \boldsymbol{\beta}'\boldsymbol{x}_j - \beta_p p_j + \varepsilon_j, \tag{1}$$

$\boldsymbol{x}_j$ is a $k \times 1$ vector of attributes of the product, including the feature that requires valuation. For mathematical convenience, it is assumed that the error terms have an extreme-value type I or Gumbel distribution. Consumers are assumed to know the realization of the error term and simply choose the alternative with maximum total utility.

To derive the standard choice model, we must calculate the probability that

[3] Unlike standard regression models, a characteristics model does not have an intercept.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 94 of 122

the jth alternative has the maximum utility, employing the assumption that the error terms have a Gumbel distribution with a scale parameter of 1. That is, since we do not observe the error terms but only the deterministic portion of utility, there is an entire region of possible errors that are consistent with the choice of a specific alternative. To derive the choice probability, we have to integrate over this region of possible values of the errors using the joint distribution of the errors. The fact that the error terms have positive support (can take on, in theory, any value in the interval $(-\infty, \infty)$) means that any choice alternative is possible even if the deterministic portion of utility for that choice alternative is small relative to all other alternatives. The Gumbel distribution produces a standard logit model for the choice of products:

$$\Pr(j) = \frac{\exp(\boldsymbol{\beta}'\boldsymbol{x}_j - \beta_p p_j)}{\sum_{j=1}^{J} \exp(\boldsymbol{\beta}'\boldsymbol{x}_j - \beta_p p_j)}. \tag{2}$$

It should be noted that utility is measured only on an interval scale with an arbitrary origin. That is, the same number can be added to the utility of all J alternatives without altering which alternative has maximum utility or the choice. This property is nothing more than the statement that only relative utility matters in choice models. We can arbitrary assign one of the products as the base alternative and express utility relative to this base alternative. In most situations, it is reasonable to assign the outside option as the base alternative. That is, one of the possibilities is that the consumer decides not to purchase one of the J products and spends his or her money on other types of goods. For example, not everyone has a point-and-shoot digital camera. This is because the base utility for nonpurchase is higher for some than the utility they would obtain net of price from purchasing in the product category. It is common, therefore, to assign a utility of 0 to the outside option, and, therefore, the utility for each of the J products is expressed relative to the outside option of nonpurchase. Thus, a model with the outside option, or possibility of nonpurchase of any of the J products, can be written as

$$\begin{aligned} \Pr(j) &= \frac{\exp(\boldsymbol{\beta}'\boldsymbol{x}_j - \beta_p p_j)}{1 + \sum_{j=1}^{J} \exp(\boldsymbol{\beta}'\boldsymbol{x}_j - \beta_p p_j)}; \\ \Pr(\text{No Purchase}) &= \frac{1}{1 + \sum_{j=1}^{J} \exp(\boldsymbol{\beta}'\boldsymbol{x}_j - \beta_p p_j)}. \end{aligned} \tag{3}$$

Of course, most firms do not observe the choice decisions of individual consumers. Instead, firms face the aggregate demand for their products, which is the sum of the quantities demanded by each possible consumer. In most product markets, there is a very large number of potential customers. It is well documented that consumers differ dramatically (Allenby and Rossi 1999) in their preferences. In the context of our choice model, this means that consumers have very different utility weights β. For example, some consumers are very price

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 95 of 122

sensitive and regard all digital cameras as roughly equivalent in terms of performance and features. For these consumers, choice is a matter of locating the lowest priced alternative. Others may be less price sensitive and more loyal to specific brands. In terms of product features, some consumers assess little or no value to the patented feature, while others may accord it substantial value. The best way to think about this is that there is a continuum of preferences for any one characteristic rather than only a small number of segments, or types of consumers. In this situation, the consumer population is best thought of in terms of the distribution of preferences. The fact that there is a very large number of different types of consumers (each with different preferences) means that we can think of there being a continuous distribution of consumer preferences. This continuous distribution is represented by a density, $p(\boldsymbol{\beta}, \beta_p)$. To sum over all possible consumers to create aggregate demand is the same as averaging the choice probabilities with respect to the density of consumer preferences. In the equation

$$\text{Sales}_j = M \times \int \Pr(j|\boldsymbol{\beta}, \beta_p)\boldsymbol{p}(\boldsymbol{\beta}, \beta_p)d\beta d\beta_p, \tag{4}$$

M is the total size of the market (the number of total potential customers). The choice probability average over the distribution of preferences is the market share that the firm can expect to attain given the configuration of the product characteristics and price. In this model in which consumers buy at most only one product, market share is equivalent to unit sales.

Product features have value to the extent to which consumers obtain utility from them. For example, if the first product characteristic is a patented product feature, then demand for that feature depends on the distribution of the first element of the β vector over the population of potential customers. If most consumers place positive and substantial weight on the patented feature, then the firm can increase sales by adding the feature while keeping the price constant or maintain sales with a higher price. Economic valuation of any product feature will require estimation of the demand system given by equations (3) and (4).

3.2. Estimating Demand

We have seen that valuation of a product feature must be driven on the demand side by the utility weight accorded that feature and the distribution of these utility weights across individuals. Thus, a critical, but not the only, component of patent valuation is to estimate the demand system. In particular, we must estimate the distribution of utility weights for the patented product feature, which requires observing demand for products with and without this particular feature. We should start by delineating the set of features that determine demand for this type of product or product category. This list might be very extensive. For example, new versions of a smartphone may incorporate changes to over 1,000 different features, only some of which we are interested in valuing for patent

purposes. If we could specify a list of the product features or characteristics, we would then, in an ideal world, conduct controlled experiments in which product features and prices are varied according to a randomized orthogonal design. We would also want to observe individual demand or choice decisions. With this panel data structure (multiple observations for the same consumer), we could estimate the utility weights for a sample of consumers. This would provide an estimate of the distribution of preference parameters and allow us to construct an estimate of aggregate demand. From this aggregate demand estimate, we could then undertake some type of equilibrium analysis to compare firm profits for products with and without the patented feature.

From a practical point of view, the estimation of demand for patented features may require data not available to us. Typically, we have access only to aggregate sales or market share data for a small number of time periods. With a great deal of aggregate share data, it is possible, in principle, to estimate the distribution of utility weights, but with only small number of observations, these estimates are apt to be unreliable. More important, we typically do not have access to data in which the same product (in terms of brand and other relevant characteristics) is sold both with and without the patented feature. Without variation in the inclusion of the patented feature, no amount of data, at either the aggregate or consumer level, will allow us to estimate the utility weights accorded the patented feature.

In some cases, we can observe aggregate sales of a product in a time series in which the feature is added at some point during the time series. Again, we might consider using this time-series variation as a way of identifying the value of the patented feature. However, such analysis must also control for other changes in the market such as the feature composition of competing brands. Moreover, for many consumer products, there is very little time-series variation in prices. Without variation in prices, we will not be able to determine how consumers might trade off the utility they obtain for the patented feature against an increase in price. For example, consider the addition of fourth-generation (4G) technology to smartphones. There is a times series of aggregate sales for smartphones both before and after the introduction of 4G technology. However, the prices of smartphones do not vary much over time, which makes it difficult to determine what sort of price premium the addition of 4G could command.

In observational data, we also face the problem of omitted product characteristics and price endogeneity (see, for example, Berry, Levinsohn, and Pakes 1995). No matter what set of characteristics we are able to measure and add to our demand model, one can always make the argument that there are omitted or unmeasured product characteristics that drive demand. If this is the case, then we might expect that firms set prices, in part, with reference to these omitted characteristics. This means that prices are not exogenous, and typically the utility weight on price β_p would be biased downward. In general, there may be unobserved drivers of demand that are correlated with price, and thus we cannot

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 97 of 122

use all of the variation in price (however small in the first place) to estimate the price coefficient.

In summary, standard observational data are inadequate for valuation of patented features for four reasons: we often do not observe the same product with and without the patented feature, we have only a short time series of aggregate data, there is often little or no price variation, and what price variation we observe may be confounded with unobserved or unmeasured product characteristics, which results in endogeneity biases.

To overcome the shortcomings of observational data, the only real option would be to collect experimental demand data at the consumer level. This might appear impractical in the context of patent litigation. It certainly would require time and a level of resources typically not available in the litigation process. Instead, we can employ a survey-based experimental method called conjoint analysis to conduct what amounts to controlled, but simulated, experiments at the consumer level. With modern Bayesian statistical methods, these data can be used to obtain estimates of the distribution of consumer preferences and aggregate demand that can then be used to compute a profit-based measure of product value.

3.3. Conjoint Surveys and Demand Estimation

While it may be impractical to conduct field experiments in which patented features are added and removed from products to assess demand, it is possible to devise a survey-based simulation that achieves the same end. In a choice-based conjoint survey, products are described as bundles of characteristics. A survey respondent is recruited and confronted with a sequence of choice tasks. In each task, the respondent is asked to choose from a menu of hypothetical products described by their characteristics. Figure 1 shows a typical choice task from a survey designed to measure demand for digital cameras. The respondent is asked to choose between four hypothetical cameras. Each hypothetical camera is described by seven characteristics. Each characteristic has a number of values or levels. For example, price has four possibilities spread over the range from $79 to $379, while there are only two possibilities for the swivel-screen feature (present or not). In addition, the respondent is given the option of electing not to purchase any of the products offered in each choice task screen.

The respondent is faced with a small number of choice tasks (in this survey there were 16 choice tasks, or screens). The conjoint choice task is designed to simulate a marketplace in which products are described by their features. In making choices, respondents reveal their preferences for each of the product characteristics (including the patented feature) in the same way as they would in the marketplace. Thus, conjoint surveys are designed to measure preferences in the spirit of revealed-preference theory. That is, we are not asking each respondent directly, what is the maximum you would be willing to pay for this camera, or how important is feature A relative to feature B? Instead, we deduce

Scenario 1 of 16	Camera 1	Camera 2	Camera 3	Camera 4	
Brand	**Canon Powershot**	**Panasonic Lumix**	**Sony Cyber-shot**	**Nikon COOLPIX**	**None of these**
Megapixels	16	16	16	16	
Optical Zoom	10x	4x	10x	4x	
Video	Full HD Video (1080p) with Stereo Microphone	HD Video (720p)	HD Video (720p)	Full HD Video (1080p) with Stereo Microphone	
Swivel Screen	No	Yes	No	Yes	
Wifi	Yes	Yes	No	Yes	
Price	$79	$279	$379	$179	
Which of these digital cameras do you prefer?	●	●	●	●	●

As a reminder, you can hover over the features to view the definitions.

Figure 1. Conjoint survey choice task screen

what these preferences are via simulated choice behavior. The assumption is that the preferences that respondents reveal in the conjoint exercise are the same as the preferences that dictate behavior in the marketplace. There is a long history of successful application of conjoint methodology and surveys in forecasting demand for new products and customer response to price changes (see Orme [2009] for many examples).

3.3.1. Conjoint Design

To undertake a rigorous conjoint survey for the purpose of estimating demand for patented features, specific criteria must be met. A representative sample must be used, major product features must be included in the design, product features or characteristics must be described in terms meaningful to the respondents, a proper randomized experimental design must be used to select the combinations of characteristics and levels used in the conjoint survey, the outside or no-choice option must be included, and the major set of competitive brands must be included.

3.3.2. Sampling Procedure

In any survey, we are extrapolating or projecting from the sample to the relevant population. In the application of conjoint surveys to patent valuation, the relevant population is the population of potential customers for the product category. The best way to ensure representativeness for this population is to start with a representative sample of all consumers in the United States and then screen this sample to those who are in the market for the products under consideration. The only sampling procedures that can guarantee representativeness are random-sampling procedures. For example, if we have a list of all U.S. residents, we would select respondents at random (each respondent is equally likely) for the screening portion of the survey.

Unfortunately, much survey work today is done with Internet panels that are not random samples. An Internet panel is a group of potential respondents who have indicated a willingness to take surveys using Web-based methods. Invitation e-mail messages with a link to the survey are sent to a random sample of the Internet panel. The problem with most Internet panel providers is that the Internet panelists are harvested via display ads and e-mail lists. This is what statisticians call a convenience sample, in that the sample is collected by any means necessary and is not designed to be representative of any population. The fact that only a random sample of the entire Internet panel is invited does not mean that the final sample of screened respondents is a random sample of the population of potential customers. Yeager et al. (2011) document that nonrandom Internet panels can yield biased estimates of population quantities. There are Internet providers who attempt a close approximation to random samples. If at all possible, higher-quality providers should be used.

If the survey sample is not obtained via high-quality, screened random samples,

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 100 of 122

then it is still possible to use the sample but only if checks for representativeness are made. The typical check for representativeness is to compare the demographic characteristics of the prescreened sample (this requires asking demographic questions prior to screening) with census-based demographics. This assures representativeness only if the product preferences of consumers are highly correlated with demographics. For example, we could find that our sample collected by nonscientific methods is representative on the basis of the gender demographic variable. That is, the sample proportion of women is the same as the population proportion. This does not, however, guarantee that the views of our sample with respect to a specific product feature are representative of the relevant population. This would be true only if preferences for the product feature are closely related to the gender of the consumer. Clearly, for most products, demographics cannot explain very much of the variation in brand and feature preferences. Our recommendation is that variables more closely related to the product category be used. For example, if we were doing a survey of smartphones, we might insert questions about ownership of smartphones by make or model and compare the market shares of our survey with those known in the U.S. market.

Considerations of sample representativeness are critical to the reliability and generalizability of any survey, conjoint or otherwise. No survey evidence should be considered admissible or relevant unless evidence of representativeness is provided.

3.3.3. Inclusion of Product Features

The heart of the demand for product features is a specification of the relevant product characteristics. For many products, the number of features is so large that it would seem impractical to ever attempt to apportion the utility or demand that can be ascribed to any one feature. One reaction is that this makes any characteristics approach to demand impractical. However, the logit choice model of demand in equation (1) has an important property called the irrelevance of irrelevant alternatives (IIA) (see, for example, Train 2003). This property implies that any set of characteristics that are constant across choice alternatives drop out of the choice probabilities. For example, partition the characteristics vector into two parts, $(\boldsymbol{x}_0, \boldsymbol{x}_1)$. The vector segment $\boldsymbol{x}_0$ varies across choice alternatives while the vector segment x_1 does not:

$$\Pr(j) = \frac{\exp(\boldsymbol{\beta}_0' \boldsymbol{x}_{0,j} + \boldsymbol{\beta}_1' \boldsymbol{x}_1)}{\sum_j \exp(\boldsymbol{\beta}_0' \boldsymbol{x}_{0,j} + \boldsymbol{\beta}_1' \boldsymbol{x}_1)} = \frac{\exp(\boldsymbol{\beta}_0' \boldsymbol{x}_{0,j})}{\sum_j \exp(\boldsymbol{\beta}_0' \boldsymbol{x}_{0,j})}. \tag{5}$$

The IIA property of a logit model[4] means that if consumers assume that only a subset of the characteristics are varied while all other characteristics are constant across choice alternatives, then we can construct valid demand estimates by testing

[4] It should be noted that although we are assuming that the irrelevance of irrelevant alternatives holds at the individual consumer level, this does not mean that it holds as a property of the aggregate demand system.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 101 of 122

only a subset of relevant product characteristics. This greatly enhances the power of a choice-based conjoint survey. In theory, as long as we assume that the logit model of demand is correct, we have to examine only a subset of product features. We will have to instruct respondents to assume, in making their choices, that all other features are constant. This is common in conjoint surveys.

In contrast, aggregate demand modeling with observational data does not hold the unobserved characteristics constant across products. If we do not observe important product features, then we might find that prices are determined, in part, by the unobserved features, which creates an endogeneity bias problem. We would have to find variables, called instruments, that move prices but are uncorrelated with the unobserved characteristics. It can be difficult to identify such variables. In theory, this is not a problem in a conjoint survey because the respondents are specifically instructed to assume that all features or characteristics other than those varied in the survey are constant across alternatives.

Taken literally, the IIA property means that we could conduct a conjoint survey with only two characteristics—the patented feature and price. In this extreme case, we are constructing a very unrealistic simulation of the marketplace.[5] For example, if one of the hypothetical products in the choice task has a very high price, it is unlikely that the respondents will hold constant other features. Their natural inclination is to assume that, perhaps, the high price indicates that this hypothetical product has very important features that are missing from the alternatives. This violation of the hold-constant instruction is much less likely if other important features are included in the conjoint survey design. This means that, even though we may wish to test only a small number of patented features, we must include many of the other important features in the product. This, of course, does not mean that we have to include all features of the product, a typical criticism of conjoint methods. All scientific models are abstractions that attempt to capture the important aspects of the problem. For example, when *Consumer Reports* provides comparisons of cars, smartphones, televisions, or other consumer products, it does not list all features but, instead, concentrates on the important features. It is important to undertake research prior to the conjoint survey design to determine the major and most important features of the product.

There are cases in which conjoint designs have been used that test only patented product features and do not include other important product features. Not only does this invite the respondents to make attributions that are not correct, but it also calls attention to the patented features. This practice can lead to an overstatement of feature value.

3.3.4. Description of Features

A general principle of conjoint survey design is that the product characteristics must be defined in terms that are understandable and meaningful to the re-

[5] As Orme (2009, p. 91) puts it, "realism begets better data."

spondents. For example, smartphone battery capacity should be specified in terms of battery life in use rather than in units of capacity such as milliamperes. This is a particularly challenging problem with patented features. Patents often describe a very specific apparatus or, in the case of a method patent, a specific set of steps that must be undertaken to practice the patent. For example, it might be impossible to obtain a patent for a generic functionality such as transmitting pictures in e-mail messages, but there may be patents that describe a specific method for attaching a picture file to an e-mail message. A conjoint survey that construes this patented feature as simply the ability to send pictures via e-mail messages would overstate the value of the patent. It may well be the case that it is possible to send pictures via e-mail messages using a method that does not infringe on the specific patent in question. If this noninfringing alternative provides the same functionality to the user, then there is no point in doing a conjoint survey. However, if alternatives to the patent provide inferior functionality, then the conjoint survey must be designed to value this specific functionality. For example, if the patented technology provides greater speed or reliability of transmission, then the challenge in the conjoint survey is to describe diminished speed or reliability to the survey respondent.

For patent litigation, it is important that the conjoint survey relate as closely as possible to the actual specifications or construals of the patent. If there are noninfringing alternatives, then the conjoint survey must be designed to test the relative value of the patented feature to the noninfringing alternatives. Features and functionality must be described to the respondent in simple and meaningful terms. This may involve the use of graphics and video descriptions.

There is no question that many patents are complex, and there will still be much debate as to the appropriate construal of the patent. We do not offer a solution to this problem, which will certainly continue to be much debated in the litigation process. However, our main point is that the valuation of patents depends critically on the valuation of product features. In order to construct a meaningful survey to value these features, descriptions of the features must be understandable by the survey respondents. This certainly means that careful pretesting will be required (for a discussion on surveys and pretesting, see Sheatsley 1983).

3.3.5. Experimental Design

Conjoint surveys are properly viewed as experiments in which the products considered in the simulated choices are designed for maximum discrimination between features. Once a set of characteristics or features (called attributes in the conjoint literature) are selected, then the levels of these attributes must be selected. For example, if we include megapixels as a characteristic, or attribute, of digital cameras, then we need to select the specific values (for example, 10 or 16 mp). Given the set of attributes and the possible levels for each attribute, the conjoint design task then consists of creating hypothetical products as bundles

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 103 of 122

of these attributes and specific levels. While this is a highly technical subject (see, for example, Box and Draper 1987, chs. 4 and 5), the central intuition is that we must vary each attribute independently of the other attributes in order to learn about the utility weights for each attribute. For example, we cannot always have those products with the patented feature be priced higher than those hypothetical products without the feature.

There are well-known and reliable ways of automatically generating choice tasks that will yield maximally informative and balanced choice task designs. It is this experimental design that makes conjoint surveys especially valuable, as these designs create hypothetical products with configurations for the purpose of revealing survey respondents' preferences. In the marketplace, we tend to see less independent variation of product features and price.

3.3.6. The Outside Option

In the real marketplace, not all potential customers purchase one of the available products. For example, although the market penetration of digital cameras is high, not everyone has one. This means that, in order to be realistic, the conjoint study must include the outside option, "none of the above," to allow respondents to opt out of the purchase. This is especially important in the evaluation of new product features. Firms invest in the development of new patented features in order to compete for customers not only with existing products but also with the hope of attracting new customers into the market. For example, the design features incorporated into the first iPad greatly expanded the market for tablet computers. If the outside option is not included in the conjoint study, then demand can come only at the expense of competing products with no growth in the overall market.

Practitioners of conjoint surveys have found that how the outside option is included makes a difference. Just adding a column for "none of the above" has been found to cause respondents to overstate their purchase intentions (Brazell et al. 2006). One possible explanation for this overly optimistic reported behavior is that respondents do not pay sufficient attention to the price attribute. Another is that respondents sometimes feel awkward rejecting products in which they believe the conjoint survey designer has a personal stake. It is common to use a dual-response mode of incorporating the outside option. In the choice task, respondents are forced to choose one option that is their preferred option. Then the respondents are asked explicitly if they would purchase their preferred product at the stated price.

3.3.7. The Set of Competing Brands

Our method of valuing a patented feature is to compute the incremental profits that the patent holder would earn from including the patented feature in the product. This depends on the structure of competition in the market. Competition is defined both by the number of competitors and by the position of

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 104 of 122

their products in the marketplace. Unless competing brands are included in the conjoint analysis, it will be impossible to make realistic estimates of the profits that can be realized from the patented feature.

This point seems to be lost on some involved in patent damages calculations. If a firm is accused of infringement, there is the natural inclination to focus on the value of the feature for the products accused of infringement. The evolution of thought in patent damages is that not all sales of the accused infringing products are due entirely to the contribution of the infringing feature. For this reason, the emphasis has been on decomposing the value of the feature in terms of what incremental sales are attributed to the infringement. This has contributed to the attraction of conjoint methods, which offer the possibility of decomposing demand. However, it is clear that if the infringing products were removed from the market, the patent holder's products that practice the patent would garner only some of the sales of the infringing product. Furthermore, the prices in the market would be expected to adjust to the withdrawal of the infringing products. Without a realistic demand system for all major brands in the market, it is not possible to calculate a meaningful new industry equilibrium for the world without the infringing products.

In summary, conjoint surveys when properly designed, pretested, and applied to representative samples can be used to estimate industry demand for patented product features. However, a proper valuation of patented features does not end with the production of conjoint data. These data must be analyzed to produce reliable estimates of aggregate demand for the relevant firms, and we must undertake equilibrium calculations.

4. Equilibrium Calculations

The value of a patent is ultimately derived from the profits that accrue to firms who practice the patent by developing products that utilize the patented technology. In fact, the standard economic argument for allowing patent holders to sell their patents is that patents will eventually find their way into the hands of those firms who can best utilize the technology to maximize demand and profits. We believe that a reasonable measure of the economic value of feature enhancement is the incremental profits that the patented feature will generate. In the equation

$$\Delta\pi = \pi(p^{\mathrm{eq}}, m^{\mathrm{eq}}|A^{*}) - \pi(p^{\mathrm{eq}}, m^{\mathrm{eq}}|A), \tag{6}$$

π is the profits associated with the industry equilibrium prices and shares given a particular set of competing products, which is represented by the choice set defined by the attribute matrix; A^{*} refers to the set of competing products, one or more of which has been enhanced by the addition of the patented feature; and A refers to the set of products without feature enhancement. This definition allows for both price and share adjustment as a result of feature enhancement.

In order to compute incremental profits as in equation (6), we must make

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 105 of 122

assumptions about the demand for product features, costs, and competitive structure. Finally, we use the standard Nash price equilibrium concept to compute equilibrium prices. We assume the following:

1. a standard heterogenous logit demand specification that is linear in the attributes (including price),
2. a cost specification with constant marginal cost,
3. single-product firms,
4. no exit (firms cannot exit or enter the market after product enhancement takes place), and
5. static Nash price competition.

Assumptions 2 and 3 can be easily relaxed. Assumption 1 can be replaced by any valid or integrable demand system. Assumptions 4 and 5 cannot be removed without imparting considerable complexity to the equilibrium computations.

The standard static Nash equilibrium in a market for differentiated products is a set of prices that simultaneously satisfy all firms' profit-maximization conditions. Each firm chooses a price to maximize its profits, given the prices of all other firms. These conditional demand curves are sometimes called the best response of the firm to the prices of other firms. An equilibrium, if it exists,[6] is a set of prices that are simultaneously the best response or profit maximizing for each firm given the others.

In a choice setting, firm demand is

$$\pi(p_j|p_{-j}) = M\mathbb{E}[\Pr(j|\boldsymbol{p}, A)](p_j - c_j), \tag{7}$$

where M is the size of the market, $\boldsymbol{p}$ is the vector of the prices of all J firms in the market, and c_j is the marginal cost of producing the firm's product. The expectation is taken with respect to the distribution of choice model parameters. In the logit case,[7]

$$\mathbb{E}[\Pr(j|\boldsymbol{p}, A)] = \int \frac{\exp(\beta' a_j - \beta_p p_j)}{\sum_j \exp(\beta' a_j - \beta_p p_j)} \boldsymbol{p}(\beta, \beta_p) d\beta d\beta_p. \tag{8}$$

The first-order conditions of the firm are

$$\frac{\partial \pi}{\partial p_j} = \mathbb{E}\left[\frac{\partial}{\partial p_j} \Pr(j|\boldsymbol{p}, A)\right](p_j - c_j) + \mathbb{E}[\Pr(j|\boldsymbol{p}, A)]. \tag{9}$$

The Nash equilibrium price vector is a root of the system of nonlinear equa-

[6] There is no guarantee that a Nash equilibrium exists for a heterogeneous logit demand model.

[7] We do not include a marketwide shock to demand, as we are not trying to build an empirical model of market shares. We are trying to approximate the firm problem. In a conjoint setting, we abstract from the problem of omitted characteristics, as the products we use in our market simulators are defined only in terms of known and observable characteristics. Thus, the standard interpretation of the marketwide shock is not applicable here. Another interpretation is that the marketwide shock represents some sort of marketing action by the firms (for example, advertising). Here we are directly solving the firm pricing problem and holding fixed any other marketing actions. This means that the second interpretation of the marketwide shock as stemming from some unobservable firm action is not applicable.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 106 of 122

tions that define the first-order condition for all J firms. That is, if we define

$$h(\boldsymbol{p}) = \begin{bmatrix} h_1(\boldsymbol{p}) = \dfrac{\partial \pi}{\partial p_1} \\ h_2(\boldsymbol{p}) = \dfrac{\partial \pi}{\partial p_2} \\ \vdots \\ h_J(\boldsymbol{p}) = \dfrac{\partial \pi}{\partial p_J} \end{bmatrix}, \tag{10}$$

then the equilibrium price vector $\boldsymbol{p}^*$ is a zero of the function $h(\boldsymbol{p})$. There are computational issues in both evaluating the firm first-order conditions and computing the root of the system of equations given in equation (10). We refer the reader to Allenby et al. (2014) for details.

5. Willingness to Pay and Willingness to Buy

In some recent patent cases, valuation of patents and associated product features is conducted via either WTP or WTB analysis. The concept of WTP should be defined as the monetary amount required to compensate a consumer for the loss of product value. For example, WTP for a product is the amount of extra income that must be given to the consumer so that the consumer is indifferent between purchasing the product and not purchasing the product but receiving the extra income. The concept of WTP for a patented product feature can be defined as the difference between WTP for a product with the enhanced feature and WTP for a product that is identical in every way except that the product no longer has the feature. The concept of WTB is often defined as the increase in product sales or market share that can be obtained by the addition of the feature to an existing product but holding prices constant.

As such, the WTP and WTB measures cannot be measures of the market value of a product feature,[8] as they do not directly relate to what incremental profits a firm can earn on the basis of the product feature. The WTP and WTB measures utilize only demand-side information and are independent of costs or the competitive structure of the industry. That is, WTP will be the same no matter how strong the competition is for the focal firm's customer base. The WTB measure holds prices fixed and does not allow the industry to adjust to a new pricing equilibrium.

[8] Ofek and Srinivasan (2002) define the market's value for an attribute improvement as the amount by which a firm can raise the price when the feature is added to a product and still command the same market share. This is a willingness-to-pay (WTP) measure and is not based on any market equilibrium computation.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 107 of 122

5.1. *Computing Willingness to Pay with Conjoint Data*

What is called WTP in the conjoint literature is one attempt to convert the part worth of the focal feature f to the dollar scale. Using a standard dummy variable coding, we can view the part worth of the feature as representing the increase in deterministic utility that occurs when the feature is turned on.[9] If the feature part worth is divided by the price coefficient,[10] then we have converted to the dollar scale, which is a ratio scale. Some call this a WTP for the product feature:

$$\text{WTP} \equiv \frac{\beta_f}{\beta_p}. \tag{11}$$

This WTP measure is often justified by appeal to the simple argument that this is the amount by which price could be raised and still leave the utility for choice alternative J the same as when the product feature is turned on. Others define this as a willingness to accept by giving the completely symmetric definition as the amount by which price would have to be lowered to yield the same utility in a product with the feature turned off as in a product with the feature turned on. Given the assumption of a linear utility model and a linear price term, these definitions are identical.[11] In the literature on conjoint methods (Orme 2001), WTP is sometimes defined as the amount by which the price of the feature-enhanced product can be increased and still leave its market share unchanged. In a homogeneous logit model, this is identical to expression (11).

The WTP measure is properly viewed simply as a scaling device. That is, WTP is measured in dollars and is on a ratio scale so that valid inter- and cross-respondent comparisons can be made. As such, WTP should properly be interpreted as an estimate of the change in WTP from the addition of the feature:

$$\Delta\text{WTP} = \text{WTP}_{f^*} - \text{WTP}_f. \tag{12}$$

Here WTP_{f^*} is the WTP for the product with the feature, and WTP_f is the WTP for the product without the feature. This measure of WTP is commonly used by conjoint analysis practitioners. However, this measure of WTP is not the standard measure of WTP for an enhanced choice set or a product innovation as defined in the economics literature (see Trajtenberg 1989). In a companion paper (Allenby et al. 2014), we compute the standard WTP measure, which should be interpreted as a measure of the consumer surplus generated by feature enhancement. Of course, welfare or consumer surplus measures are not the same as incremental firm profits, the relevant measure for patented feature valuation. With uniform pricing and competition, firms are not able to extract all of the

[9] For feature enhancement, a dummy coding approach would require that we use the difference in part worths associated with the enhancement in the WTP calculation.
[10] We have defined the price coefficient such that this is always positive. See Section 3.3.
[11] In practice, reference price effects often make willingness to accept differ from WTP (see Viscusi and Huber 2012), but in the standard economic model these are equivalent.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 108 of 122

consumer surplus, and the WTP measure could overstate the value of a patented feature.

5.2. Willingness to Buy

In some analyses, product features are valued using a WTB concept. The WTB is the change in market share that will occur if the feature is added to a specific product. In the expression

$$\mathrm{WTB} \equiv \mathrm{MS}(j|\boldsymbol{p},\ A^{*}) - \mathrm{MS}(j|\boldsymbol{p},\ A), \tag{13}$$

MS(j) is the market share equation for product j. The market share depends on the entire price vector and the configuration of the choice set. Expression (13) holds prices fixed as the feature is enhanced or added. The market share equations are obtained by summing the logit probabilities over possibly heterogeneous (in terms of taste parameters) customers.

The WTB assumes that firms will not alter prices in response to a change in the set of products in the marketplace as the feature is added or enhanced. In most competitive situations, if a firm enhances its product and the other competing products remain unchanged, we would expect the focal firm to be able to command a somewhat higher price, while the other firms' offerings would decline in demand, and therefore the competing firms would reduce their prices.

5.3. Why Willingness to Pay and Willingness to Buy Are Inadequate

Neither WTP nor WTB take into account equilibrium adjustments in the market as one of the products is enhanced by addition of a feature. For this reason, we cannot view WTP as what a firm can charge for a feature-enhanced product, nor can we view WTB as the market share that can be gained by feature enhancement. Computation of changes in the market equilibrium due to feature enhancement of one product is required to develop a measure of the economic value of the feature. In many cases, WTP will overstate the price premium afforded by feature enhancement, and WTB will also overstate the impact of feature enhancement on market share. Equilibrium computations in differentiated-product cases are difficult to illustrate by simple graphical means. In this section, we use standard demand and supply graphs to provide an informal intuition as to why WTP and WTB will tend to overstate the benefits of feature enhancement.

Figure 2 shows a standard industry supply-and-demand setup.[12] The demand curve is represented by the downward-sloping lines, while D denotes demand without the feature, and D* denotes demand with the feature. The vertical difference between the two demand curves is the change in WTP as the feature is added. We assume that addition of the feature may increase the marginal cost of production (note that for some features, such as those created purely via

[12] Here we consider the case of a competitive industry to provide an intuition for why the WTP and willingness-to-buy measures are not equilibrium quantities. In the case of imperfect competition or oligopoly, it is not possible to represent the equilibrium in this manner.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 109 of 122

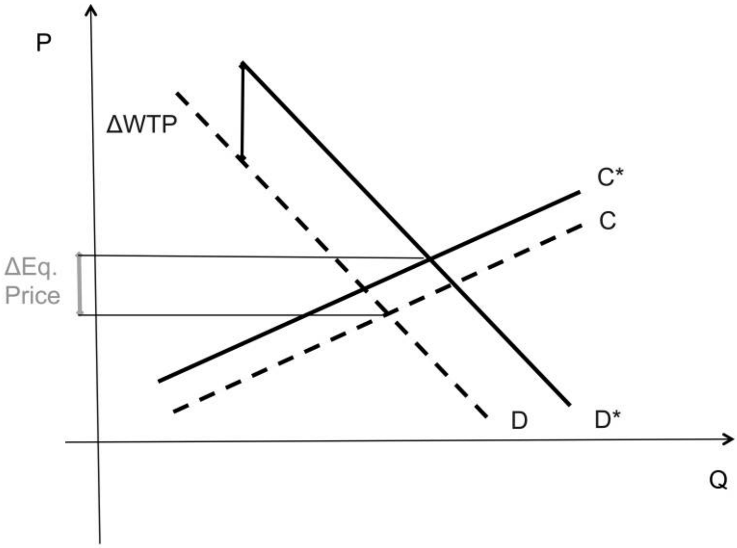

Figure 2. Difficulties with a willingness-to-pay analysis

software, the marginal cost will not change). The cost curves are marked C and C*. It is easy to see that, in this case, the change in WTP exceeds the change in equilibrium price.

The analogous situation is shown for WTB in Figure 3. We have the same cost and demand curves, but we illustrate the WTB exercise, which is to compute the change in quantity sold assuming that prices do not change. The WTB clearly overstates the changes in equilibrium quantity demanded. Figures 2 and 3 show very clearly that both WTP and WTB are purely demand-based quantities that do not take into account changes in prices and costs as the feature is enhanced and a new industry equilibrium is achieved.

5.4. Willingness to Pay in the Case of Heterogeneous Customers

Even in the case of homogeneous customers, we have seen that WTP should not be regarded as a proper measure of economic value. In the case of heterogeneous consumers, additional problems are associated with the WTP concept. In almost all choice-based conjoint settings, hierarchical Bayes methods are used to estimate the choice model parameters. In the hierarchical Bayes approach (see, for example, Rossi, Allenby, and McCulloch 2005, ch. 5 and app. A), each respondent may have different logit parameters β and β_p, and the complete posterior distribution is computed for all model parameters, including individual respondent-level parameters. The problem, then, becomes how to summarize the distribution of WTP, which is revealed via the hierarchical Bayes analysis.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 110 of 122

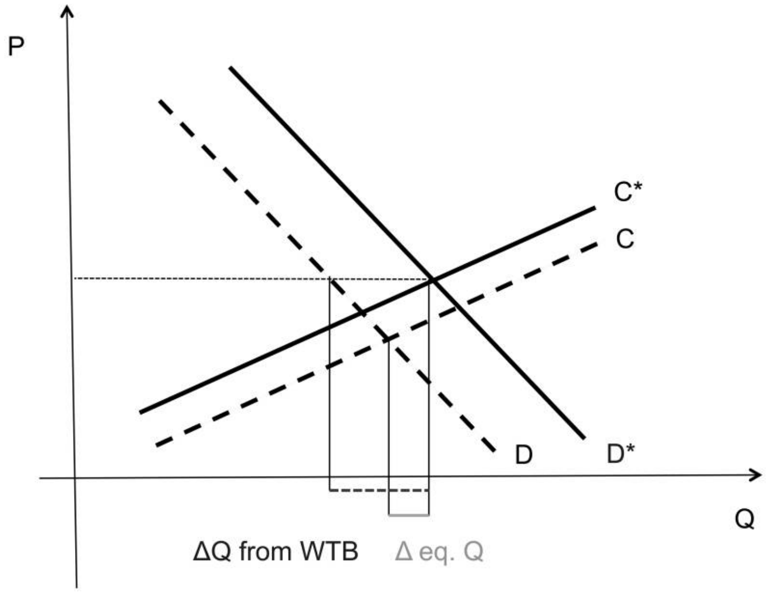

Figure 3. Difficulties with a willingness-to-buy analysis

The concept of WTP provides no guidance as to how this distribution should be summarized. One natural summary would be that the expectation is taken over the distribution of model parameters:

$$\mathbb{E}[\mathrm{WTP}] = \int \frac{\beta_f}{\beta_p} \boldsymbol{p}(\beta_f, \beta_p|\mathrm{Data}) d\beta_f d\beta_p.$$

However, there is no compelling reason to prefer the mean over any other scalar summary of the distribution of WTP. Some propose using a median value of WTP instead. Again, there are no economic arguments as to why the mean or median or any other summary should be preferred. The statistical properties of various summaries (for example, mean versus median) are irrelevant, as we are not considering the sampling performance of an estimator but rather the appropriate summary of a population distribution. A proper economic valuation will consider the entire demand curve as well as competitive and cost considerations. Equilibrium quantities will involve the entire distribution via the first-order conditions for firm profit maximization. These quantities cannot be expressed as a function of the mean, median, or any other simple set of scalar summaries of the distribution of WTP.

However, it is possible to provide a rough intuition as to why the mean of WTP may be a particularly poor summary of the distribution for equilibrium

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 111 of 122

computations. It is the marginal rather than the average consumer that drives the determination of equilibrium prices. Exactly where in the distribution of WTP the marginal customer will be is determined by nature of the distribution as well as where supply factors that slice into the distribution of WTP lie. It is possible to construct cases in which the average WTP vastly overstates the WTP of the marginal customer. If the bulk of the market has a low value of WTP, and there is a small portion of the market with an extremely high value of WTP, then a profit-maximizing firm may set the price much lower than the average WTP so as to sell to the majority of potential customers who have relatively low values of WTP. There are situations in which the greater volume from low-WTP consumers outweighs the high margins that might be earned from the high-WTP segment. In these cases, the mean WTP will vastly overstate the price premium a firm will charge over cost for a product. It is more difficult, but possible, to construct similar scenarios for the median WTP.

One of the major problems with using any measure of the central tendency of the distribution of WTP (mean, median, or mode) is that this includes consumers whose WTP is insufficient to be in the market. That is, our surveys should qualify respondents to be in the market for the products (with a screening question such as "Do you plan to buy a digital camera in the next 6 months?"). However, simply averaging WTP over all survey respondents averages in those whose WTP for the product as a whole is below the market price of any product and, therefore, would not purchase in the product category. This is a downward bias for WTP computations. Instead of debating whether a particular WTP computation is upward or downward biased for feature valuation, our view is that we should employ an incremental-profits measure of economic value for the patent context.

6. An Illustration Using the Digital Camera Market

To illustrate our proposed method for patent valuation and to contrast our method with standard WTP and WTB methods, we consider the example of the digital camera market. We designed a conjoint survey to estimate the demand for features in the point-and-shoot submarket. This is a consumer electronics category that is very similar to the product categories in which there have been patent litigation (smartphones, tablets, and gaming consoles).

We considered the following features with associated levels: (1) brand: Canon, Sony, Nikon, and Panasonic; (2) pixels: 10 and 16 mp; (3) zoom: 4 × and 10 × optical; (4) video: high-definition (720p) and full high-definition (1080p) with microphone; (5) swivel screen: no or yes; (6) Wi-Fi: no or yes; (7) price: $79–$279. We chose the swivel-screen feature as the focus of our analysis. This feature

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 112 of 122

Figure 4. Swivel-screen attribute

is illustrated in Figure 4. We consider the problem of valuing a patent that covers the swivel-screen feature.[13]

The conjoint design was a standard fractional factorial design in which each respondent viewed 16 choice sets, each of which featured four hypothetical products. A dual response mode was used to incorporate the outside option. Respondents were first asked which of the four profiles presented in each choice task was most preferred. Then respondents were asked if they would buy the preferred profile at the stated price. If the response was no, then this response was recorded as the outside option, "none of the above." Respondents were screened to include only those who owned a point-and-shoot digital camera and who considered themselves to be a major contributor to the decision to purchase this camera.

The survey was fielded to the Sampling Surveys International Internet panel in August 2013. We received 501 completed questionnaires.[14] We recorded time to complete the conjoint portion of the survey. The median time to complete was 220 seconds, or about 14 seconds per conjoint task. The 25th percentile is 151 seconds, and the 75th percentile is 333 seconds. To check sensitivity to time spent on the survey, we conducted analyses deleting the bottom quartile of the respondents and found little change. It is a common and well-accepted practice to remove respondents who "straight-line," or always select the same option (such as the left-most choice). The idea is that these straight-line respondents are not putting sufficient effort into the choice task. Of our 501 complete questionnaires, only two respondents displayed straight-line behavior and were elim-

[13] We are not aware of a specific patent related to the swivel-screen apparatus, but this feature is representative of many patented features.

[14] This study was part of a wave of four other very similar conjoint studies on digital cameras, each with the same screening criteria. For all studies in the wave, 16,185 invitations were sent to panelists, and 6,384 responded. Of those who responded to the invitation, 2,818 passed screening, and of those passing screening 2,503 completed the questionnaire. Thus, the overall completion rate is 89 percent, which is good by survey standards.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 113 of 122

inated. We also eliminated six respondents who always selected the same brand and two respondents who always selected the high-price brand. Our reasoning is that these respondents did not appear to be taking the trade-offs conjoint exercise seriously. We also eliminated 23 respondents who always selected the outside option, as their part worths are not identifiable without prior information. Thus, our final sample size is 468.

Our conjoint survey data consist of repeated choice observations for a set of respondents. This allows for the possibility that each of the logit coefficients in the choice model is different for each survey respondent. We use a Bayesian hierarchical logit model that specifies a normal distribution of coefficients (preferences) (see, for example, Rossi, Allenby, and McCulloch 2005). This can be interpreted as a Bayesian analysis of a random coefficient logit model. In the Bayesian approach, priors are placed on the normal distribution parameters. This model can be represented by a sequence of three distributions:

$$\begin{aligned} y_i &\sim \text{Multinomial}(\boldsymbol{\pi}(\boldsymbol{\beta}_i | \boldsymbol{p},\ A)); \\ \boldsymbol{\beta}_i &\sim N(\boldsymbol{\mu}_\beta,\ V_\beta); \\ \boldsymbol{\mu}_\beta,\ V_\beta &\sim N(\bar{\boldsymbol{\mu}},\ a_\mu^{-1} V_\beta) \times \text{IW}(\nu,\ V). \end{aligned} \tag{14}$$

The first distribution is a multinomial choice outcome, where $\boldsymbol{\pi}$ is a vector of choice probabilities, one for each product in the market. The choice probabilities are driven by choice-alternative characteristics as in equation (1) as well as individual-specific utility weights β_i. The second distribution is the random coefficient model, and the third represents priors on what Bayesians call the hyperparameters of the normal distribution. The Bayesian hierarchical approach overcomes many of the limitations of standard maximum likelihood estimation of random coefficient choice models. For our purposes, it is especially important to recognize that we are not able to know the parameters of the distribution of preferences in the population precisely after observing our sample of about 500 respondents. In a Bayesian procedure, this uncertainly is characterized by the posterior distribution of the normal distribution parameters (μ, V_β). Thus, we must integrate out or average over possible values of these distributional parameters to form what is known as the posterior predictive distribution:

$$\boldsymbol{p}(\boldsymbol{\beta}|\text{Data}) = \int \varphi(\boldsymbol{\beta}|\boldsymbol{\mu},\ V_\beta)\boldsymbol{p}(\boldsymbol{\mu},\ V_\beta|\text{Data})d\boldsymbol{\mu}dV_\beta. \tag{15}$$

Equation (15) reflects the influence of both the model (the normal random coefficient distribution) and the data through the posterior distribution of the normal hyperparameters. The resulting distributions will be symmetric but of fatter tails than in the normal distribution. We also impose the constraint that the price coefficient is strictly negative (for details, see Allenby et al. 2014).

We find aggregate demand by taking the expectation of choice probabilities with respect to the distribution of preference parameters over the population.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 114 of 122

Table 1
Posterior Mean of Aggregate Demand Elasticities

	$\text{MktShare}_{\text{Nikon}}$	$\text{MktShare}_{\text{Canon}}$	$\text{MktShare}_{\text{Sony}}$	$\text{MktShare}_{\text{Panasonic}}$
p_{Nikon}	−1.60	.39	.41	.28
p_{Canon}	.36	−1.71	.50	.33
p_{Sony}	.39	.52	−1.82	.43
$p_{\text{Panasonic}}$	.34	.43	.56	−2.14

The distribution of the part worths is the correct predictive distribution of preferences. In order to shed some light on the substitution structures found in aggregate demand, we compute the posterior distribution of the market share elasticity matrix:

$$\frac{\partial \mathrm{MS}(i)}{\partial lnp_j} = \frac{\partial}{\partial lnp_j} \int \Pr\,(i|\boldsymbol{\beta})\boldsymbol{p}(\boldsymbol{\beta}|\mu,\ V_{\text{Beta}})d\beta. \tag{16}$$

We note that the elasticities in equation (16) are functions of the normal distribution hyperparameters. We average the elasticities with respect to draws from the posterior distribution of the hyperparameter to obtain the posterior mean of these elasticities and present them in Table 1. These own-price elasticities are quite reasonable and imply a reasonably high markup of about three times the cost. The cross-price elasticities are large but significantly smaller than the own-price elasticities. This means that each of these brands (products) is not regarded as very substitutable for other digital cameras, but there is a good deal of price sensitivity for the category purchase decision. This comes from heterogeneity in brand preferences across consumers. That is, there are some consumers, for example, who prefer the Nikon brand. This means that Nikon cameras are not regarded as close substitutes for other brands. This does not mean, however, that those people who prefer Nikon are not also price sensitive, which drives up the own-price elasticity for the Nikon-branded product. Strong brand preferences are common in consumer goods markets, so we regard this industry elasticity structure as fairly typical.

6.1. Changes in Equilibrium Prices and Shares

We have argued that the economic value of a patented feature should be the incremental profits that accrue to the firm that practices the patent by incorporating the patented feature. Before showing changes in profits, we first explore changes in equilibrium prices and associated market shares. We compute equilibrium prices with and without the patented feature to provide an idea of how much of a price premium a firm can charge and how market shares will adjust in the new industry equilibrium.

Unlike the WTP computations, any equilibrium calculation will depend, as it should, on which brand and product configuration is being enhanced by addition of the patented feature. Here we consider the change in equilibrium outcomes

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 115 of 122

Table 2
Posterior Means of Equilibrium Prices

	Nikon	Canon	Sony	Panasonic
Without swivel screen	207.59	188.79	172.87	185.00
With swivel screen	263.83	186.55	172.24	177.84
Change	56.24	−2.24	−.63	−7.16

Table 3
Equilibrium Market Shares

	Nikon	Canon	Sony	Pana-sonic	Outside Good
Without swivel screen	10.7	12.7	13.3	10.3	53.0
With swivel screen	10.9	12.8	13.3	10.5	52.5

from adding the swivel-screen display to the Nikon base product (a Nikon brand camera with all attributes set at their lowest value except, of course, price, which is not constrained). The value conferred by the addition of the swivel screen will also depend on the configuration of other competing products. For illustration purposes only, we considered a competitive set that consists of three other brands (Canon, Nikon, and Panasonic) all similarly configured at the base level of attributes without the swivel-screen feature. We set the marginal cost for all brands at $75. When the swivel-screen feature is added, marginal cost is increased by $5 to $80.

Table 2 presents the posterior means of the equilibrium prices computed with and without the swivel-screen addition to the Nikon product. Adding the swivel screen gives the Nikon brand more effective market power relative to the other branded competitors that do not have the feature. Not only does Nikon find it optimal to raise the price, but the stronger competition and diminished value of the other brands force them to lower prices in equilibrium.

Table 3 displays the equilibrium market shares for each of the four branded products and the outside good calculated with and without the swivel-screen display. Only very minor share changes are observed. The Nikon product with the swivel screen gains share, primarily from the outside alternative. As we have seen, the other brands reduce their prices in equilibrium to compensate for the greater desirability of the Nikon product. We show below that these share results are much different from a WTB analysis in which prices are not allowed to adjust. A WTB analysis will overstate the share gain for Sony from feature enhancement, as the prices in a WTB analysis are not allowed to equilibrate in the new structure of competing products.

6.2. Changes in Equilibrium Profits

Our entire perspective on patent valuation is based on the incremental profits that accrue to a firm whose products include the patented feature. Obviously,

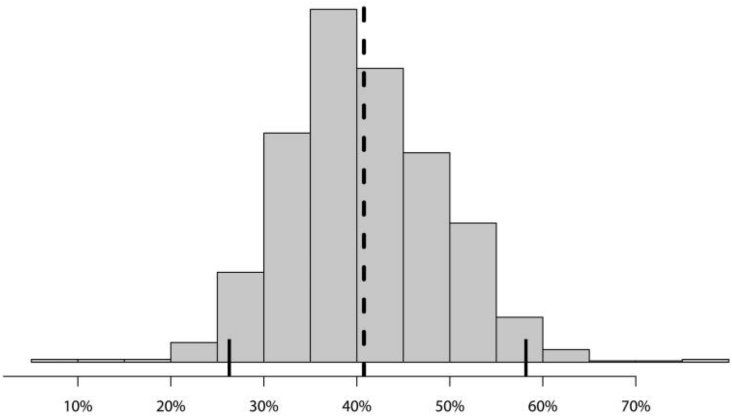

Figure 5. Posterior distribution of the change in Nikon's equilibrium profits from addition of a swivel screen.

the absolute dollar levels of profits depend on the size of the market. In order to present incremental profits in an interpretable manner, we report the posterior distribution of the percentage change in profits. That is, we again consider the addition of the swivel screen to the base Nikon product:

$$\%\Delta\pi = \frac{\pi_{\text{Nikon}}(p^{\star}_{\text{SS}}) - \pi_{\text{Nikon}}(p^{\star}_{\text{w/oSS}})}{\pi_{\text{Nikon}}(p^{\star}_{\text{w/oSS}})}.$$

To compute the posterior distribution of the percentage change in Nikon profits, we recompute the price and share equilibria for each draw from the posterior distribution of the normal distribution hyperparameters (μ, V_{β}). Figure 5 shows this posterior distribution, which is centered over an increase in equilibrium profits of about 42 percent; the posterior mean is indicated by the heavy dashed vertical line. This increase in profits is due primarily to the price premium that Nikon is now able to charge for the product. It should be emphasized that our equilibrium computations take into account the adjustments that will occur in the market after the introduction of the enhanced Nikon product. A 95 percent posterior interval (indicated by the small vertical bars along the X-axis) for percentage change in profits is quite wide, extending from about 26 percent to 58 percent. This reflects an important finding that, even with 500 respondents, we do not have a great deal of precision in the estimates of quantities relevant to patent valuation.

6.3. A Damages Point of View

The problem of calculating damages due to patent infringement can also be solved by application of our principle of economic valuation with incremental

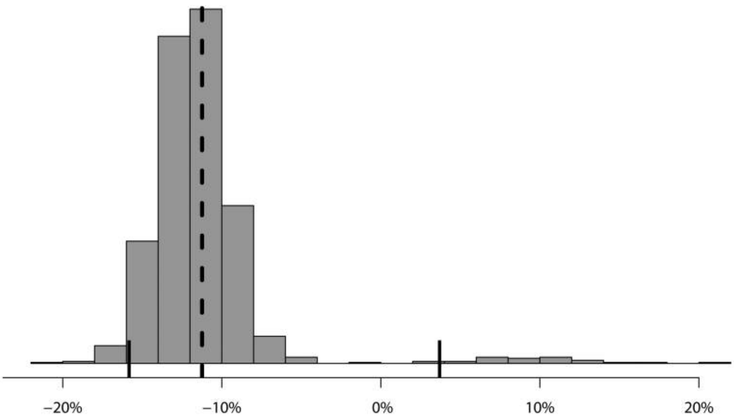

Figure 6. Posterior distribution of the change in Nikon's profits from Sony infringement

profits. Consider the case in which Nikon owns the patent for the swivel-screen technology. Sony infringes on this patent by introducing a point-and-shoot camera with the same swivel-screen feature. How has Nikon been injured by this act of infringement? The answer is clearly that Nikon is injured only to the extent that its profits are lower than in the case in which Sony does not infringe. Our demand models and equilibrium calculations allow us to compute this loss of profits. Profits in the counterfactual in which Sony does not infringe can be computed simply from the equilibrium prices and shares shown above. We then add the swivel screen to both the Nikon and Sony models and compute the associated price equilibrium. The percentage change in profits for Nikon represents the damages that can be attributed to the Sony infringement:

$$\%\Delta\pi_{\text{Nikon}} = \frac{\pi_{\text{Nikon}}(p^{\star}_{\text{Nikon, Sony SS}}) - \pi_{\text{Nikon}}(p^{\star}_{\text{Nikon Only SS}})}{\pi_{\text{Nikon}}(p^{\star}_{\text{Nikon Only SS}})}. \tag{17}$$

Here $p^{\star}$ refers to the equilibrium prices. These equilibrium computations take into account that if Sony were to infringe, Nikon might be forced to lower prices; this is often termed the problem of price erosion. In sum, this method takes full account of all ways in which Sony's infringement might harm Nikon.

Figure 6 shows how Nikon's profits would decline in the face of stiffer competition from the Sony product with the infringing feature. The posterior mean (indicated by the vertical dashed line) is an 11 percent reduction in Nikon's profits. When Sony adds the swivel screen to its product, demand for the Sony product increases dramatically. If Nikon were to keep its price at the equilibrium price that would prevail with only the Nikon product having the swivel screen, then Nikon would lose significant share to Sony. However, we would not expect Nikon to simply stand by and not react to the more formidable competition

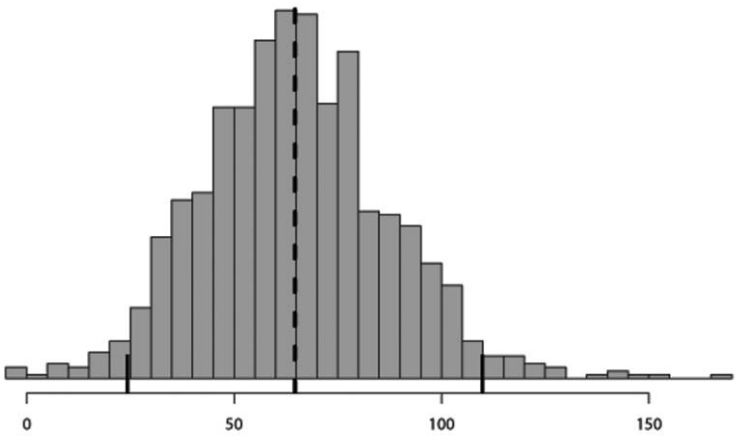

Figure 7. Posterior distribution of $\mathbb{E}$[WTP] (in dollars)

from Sony. In response to the upgrade of the Sony product, Nikon would be forced to reduce its price significantly. In the end, the new equilibrium results in Nikon reducing its price from \$263.55 to \$225.12 and maintaining approximately the same share as before the entry of the swivel-screen Sony product. This illustrates the importance of undertaking equilibrium calculations instead of simply holding prices or shares constant.

6.4. Willingness to Pay

As indicated in Section 5, the most common practice in product and feature valuation is a WTP analysis. We can easily use our digital camera data to reproduce this analysis for the swivel-screen feature. Since WTP is a consumer-specific concept, any WTP analysis for a set of consumers will have to summarize the distribution of WTP across consumers. A sensible summary would be to estimate the expected or population average WTP, $\mathbb{E}$[WTP]. The term $\mathbb{E}$[WTP] is the mean WTP across the population represented by our normal model of preferences:

$$\mathbb{E}[\mathrm{WTP}|\boldsymbol{\mu}, V_\beta] = \int \frac{\beta_f}{\beta_p} \boldsymbol{p}(\beta_f, \beta_p|\boldsymbol{\mu}, V_\beta) d\beta_f d\beta_p. \tag{18}$$

To construct the posterior distribution of $\mathbb{E}$[WTP], we use the draws from the posterior distribution of $(\boldsymbol{\mu}, V_\beta)$. For each draw of the hyperparameter, a very large number of draws are made from the normal distribution of preferences (we used 10,000) to approximate the integral in equation (18). We do this for each draw from the posterior of the hyperparameters to build up the posterior predictive distribution of the mean WTP.

Figure 7 shows the posterior distribution of WTP. This distribution is centered

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 119 of 122

over very large dollar values (mean of \$64.66 and median of \$63.71). The posterior interval for WTP extends from \$25 to \$110 (indicated by the small vertical lines), which reveals a considerable amount of uncertainty. This high level of posterior uncertainty is all the more remarkable considering the large sample size of around 500 and an orthogonal design with prices varying from \$79 to \$279. This suggests that considerably larger samples would be required to produce precise estimates of WTP. Of course, larger and more informative data would not overcome the conceptual limitations of WTP as a method of valuation. The WTP analysis provides an estimate of the expected WTP, which is larger than the equilibrium price premium reported in Table 2 and is in accordance with the intuition that this is apt to be the case when the WTP of the marginal consumer is different from the average WTP.[15]

6.5. Willingness to Buy

There are two possible WTB analyses relevant to patent valuation. The value of the patent to the firm practicing the patent (under the assumption of no infringement) might be informed by a WTB analysis in which we measure the gain in market share for a firm that adds the patented feature to an existing product. The WTB would be computed simply by enhancing the product while holding the prices constant. If we take the example of adding the swivel-screen feature to a Nikon product, then we have already seen that the primary source of value in an equilibrium analysis is not the gain in share made possible by the patented feature but rather the increased price premium. Figure 8 shows a WTB analysis for the simple case of the addition of the swivel screen to the Nikon product. The WTB analysis shows a very large increase in Nikon's market share of almost 5 percentage points. This is a misleading estimate for two reasons: Nikon will not find it profit maximizing to keep the price constant when the swivel-screen feature is added, and competing brands will adjust their prices downward. We have already seen that the change in equilibrium shares is minimal when these price adjustments are made.

The second WTB analysis that could be undertaken is one to examine the loss of Nikon's market share when Sony enters the market with an infringing product. That is, we consider the market share of the Nikon product with the swivel screen in a market in which no other competitor has the patented swivel-screen feature and compare this with the market share of the Nikon product in a market in which the Sony product also has the swivel-screen feature. All share computations hold the price of the Nikon product constant at the equilibrium price that would prevail in the market prior to Sony's infringement. In other words, Nikon does not respond by lowering its price in the face of the Sony entry. This shows the flaw in the WTB analysis; it takes a static and nonequi-

[15] While the difference between the equilibrium price premium and the average WTP is relatively small in this example, if you add the swivel-screen feature to the Sony brand, the price premium will be only about one-half of the $\mathbb{E}$[WTP] estimate.

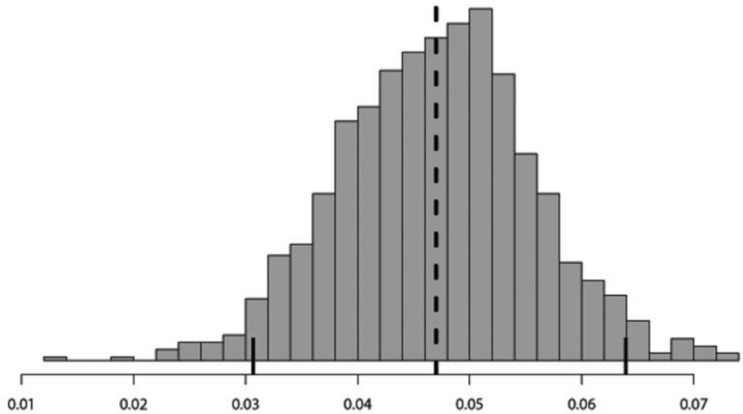

Figure 8. Posterior distribution of the change in Nikon's share from addition of the swivel-screen feature.

librium view. In recent *Apple v. Samsung* litigation, this method of WTB analysis was employed. Figure 9 shows the posterior predictive distribution of the decline in Nikon's share resulting from a WTP analysis that does not account for Nikon's price reduction in response to the entry of Sony with an infringing product. The distribution is centered on a decline in Nikon's share of about 2 percent. Equilibrium computations show no appreciable change in Nikon's share because Nikon finds it optimal to lower its price in response to Sony's infringement as detailed in Section 6.3. Thus, a damages analysis that uses the WTB estimate of lost market share would be misleading and tend to overestimate damages.

7. Conclusions

Our perspective is that a patent for a product feature has value only to the extent to which firms can earn incremental profits from the addition of the patented feature. Incremental profits are determined not only by the value that individual consumers place on the patented feature but also on the nature and extent of competition in the relevant market. This means that in order to implement our economic paradigm for patent valuation, we need to be able to estimate demand for product features and model competitive behavior. In many situations, observational data on the products, prices, and quantity demanded are not sufficient to estimate demand for patented features, and we must turn to survey-based methods.

A conjoint survey attempts to simulate marketplace purchase behavior for a sample of consumers. Product features are manipulated according to an experimental design, and the respondents are asked to choose from among hypothetical

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 121 of 122

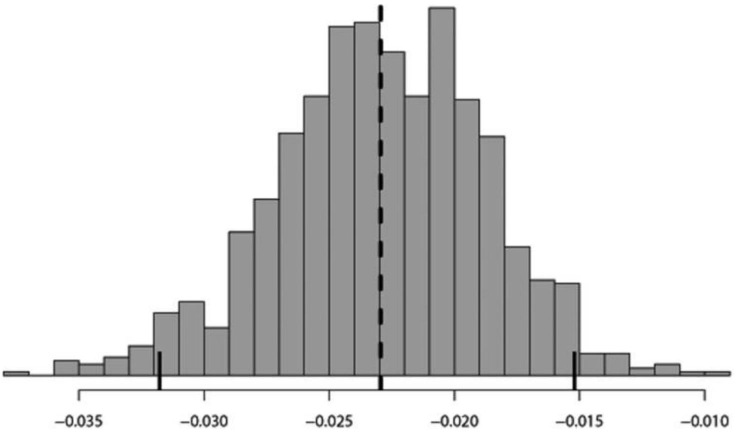

Figure 9. Posterior distribution of the change in Nikon's share when Sony infringes

products defined by product features. This eliminates many of the drawbacks of using observational data, which are frequently not sufficiently informative to be useful in patent valuation. However, conjoint studies must be designed specifically for demand and profit calculations. In particular, the set of competitive brands available in the marketplace must be included, and the choice tasks must include the outside option.

We illustrate our approach with conjoint data obtained from a survey on digital cameras. After calibrating a demand system for digital cameras that is based on a characteristics model, we compute incremental profits from adding a patented feature to one or more of the competing products. Equilibrium calculations allow prices to adjust as a feature is added to a product, which generates more realistic measures of economic value.

Finally, while we have demonstrated that conjoint data can be used to value patented features in an incremental-profits approach, the demands are greater on the design and analysis of conjoint data if the data are to be used for patent valuation. We have shown that the standard-size samples are inadequate to precisely estimate profit-based measures of patent value.

References

Allenby, Greg M., Jeff D. Brazell, John R. Howell, and Peter E. Rossi. 2014. Economic Valuation of Product Features. Working Paper No. 2013-05-01. Fisher College of Business, Ohio State University, Columbus.

Allenby, Greg M., and Peter E. Rossi. 1999. Marketing Models of Consumer Heterogeneity. *Journal of Econometrics* 89:57–78.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile Prices in Market Equilibrium. *Econometrica* 63:841–90.

Case 5:15-md-02617-LHK Document 720-21 Filed 03/10/17 Page 122 of 122

Box, George E. P., and Norman R. Draper. 1987. *Empirical Model-Building and Response Surfaces*. Hoboken, N.J.: John Wiley & Sons.

Brazell, Jeff D., Christopher G. Diener, Ekaterina Karniouchina, William L. Moore, Válerie Séverin, and Pierre-Francois Uldry. 2006. The No-Choice Option and Dual Response Choice Designs. *Marketing Letters* 17:255–68.

McFadden, Daniel L. 1983. Econometric Analysis of Qualitative Response Models. Pp. 1395–1456 in vol. 2 of *Handbook of Econometrics*, edited by Zvi Griliches and Michael D. Intriligator. Amsterdam: North-Holland.

Ofek, Elie, and V. Srinivasan. 2002. How Much Does the Market Value an Improvement in a Product Attribute? *Marketing Science* 21:398–411.

Orme, Bryan K. 2001. Assessing the Monetary Value of Attribute Levels with Conjoint Analysis: Warnings and Suggestions. Unpublished manuscript. Sawtooth Software, Sequin, Wash.

———. 2009. *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. Madison, Wis.: Research Publishers.

Petrin, Amil. 2002. Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy* 110:705–29.

Rossi, Peter, Greg Allenby, and Robert McCulloch. 2005. *Bayesian Statistics and Marketing*. Hoboken, N.J.: John Wiley & Sons.

Sheatsley, Paul. 1983. Questionnaire Construction and Item Writing. Pp. 195–230 in *Handbook of Survey Research*, edited by Peter H. Rossi, James D. Wright, and Andy B. Anderson. New York: Academic Press.

Train, Kenneth E. 2003. *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.

Trajtenberg, Manuel. 1989. The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners. *Journal of Political Economy* 97:444–79.

Viscusi, W. Kip, and Joel Huber. 2012. Reference-Dependent Valuations of Risk: Why Willingness-to-Accept Exceeds Willingness-to-Pay. *Journal of Risk and Uncertainty* 44: 19–44.

Yeager, David S., Jon A. Krosnick, Linchiat Chang, Harold S. Javitz, Matthew S. Levendusky, Alberto Simpser, and Rui Wang. 2011. Comparing the Accuracy of RDD Telephone Surveys and Internet Samples Conducted with Probability and Non-probability Samples. *Public Opinion Quarterly* 75:709–47.